Short sequence-paper

A novel root-specific gene, *MIC-3*, with increased expression in nematode-resistant cotton (*Gossypium hirsutum* L.) after root-knot nematode infection ☆

Xiang-Dong Zhang [a], Franklin E. Callahan [b,c,*], Johnie N. Jenkins [b,c], Din-Pow Ma [a], Mehmet Karaca [b], Sukumar Saha [b,c], Roy G. Creech [b]

[a]*Department of Biochemistry and Molecular Biology, Box 9650, Mississippi State University, Mississippi State, MS 39762, USA*
[b]*Department of Plant and Soil Sciences, Mississippi State University, Mississippi State, MS 39762, USA*
[c]*Crop Science Research Laboratory, USDA/ARS, Mississippi State, MS 39762, USA*

Abstract

A full-length cDNA, *MIC-3*, has been identified from a λZAPII cDNA library constructed from the mRNA of nematode-resistant cotton (*Gossypium hirsutum* L.) roots after infection with root-knot nematode (*Meloidogyne incognita*). The putative open reading frame of *MIC-3* encoded a protein of 141 amino acids with a calculated molecular mass of 15.3 kDa. Seven alternative polyadenylation sites have been identified for the *MIC-3* transcripts, and the major transcripts are the longest ones. The *MIC-3* gene contains a single intron within its coding region and belongs to a novel, multi-gene family containing up to six members. Expression of *MIC-3* is root localized and specifically enhanced in the nematode induced, immature galls of resistant cotton line M-249, suggesting that *MIC-3* may play a critical role in the resistance response to root-knot nematode. Published by Elsevier Science B.V.

Keywords: Cotton; Nematode; Host resistance; Root gene expression; Alternative polyadenylation

Root-knot nematodes (RKN) (*Meloidogyne* spp.) are obligate sedentary endoparasites and cause severe losses in many crops worldwide [1]. The second-stage nematode juveniles (J2) penetrate the root, migrate to the vascular tissue, and induce the formation of multinucleate "giant cells" as permanent feeding sites [2]. During giant cell formation, nearby cells expand and divide forming enlarged galls or root-knots. In resistant tomato plants, J2 stage RKN penetrate roots as in susceptible plants, but a localized cell death, involving a hypersensitive response, inhibits giant cell development at the site where feeding is initiated [3]. Some cotton (*Gossypium hirsutum* L.) lines exhibit high levels of RKN resistance [4–6], which is conferred by at least two unidentified genes [7–10]. The size and index of newly formed galls were not significantly different between the resistant and susceptible lines during early stages of infection (6–10 days after inoculation, DAI); however, RKN reproduction was inhibited ($\geq 98\%$) in resistant plants by the end of the normal reproductive cycle for RKN (60–80 DAI) [10]. In contrast to susceptible lines, the further development of J2 stage juveniles is arrested in the resistant cotton lines by 8–10 DAI, but little is known about the alterations in plant gene expression associated with this resistance response [10]. The genes with increased expression specifically in the resistant plant roots after nematode infection are likely to be critical for the establishment of resistance mediated by the nematode resistance gene(s), whereas genes with elevated expression in both resistant and susceptible roots might be involved in general defense mechanisms [11].

Our previous studies identified a 14-kDa protein, which was induced in the immature root galls of a RKN-resistant cotton isoline soon after inoculation (8–10 DAI) [12]. Based on limited internal amino acid sequence (MKVDGYYG) of the 14-kDa protein [12], a degenerate reverse primer,

R1 (5′-CCRTARTANCCRTCNACYTTCAT-3′, nucleotide (nt) positions 169–191 in Fig. 1) (Y = C/T, R = A/G, N = A + C + G + T), was designed to amplify the cDNA sequence encoding the 14-kDa protein. A cotton λZAPII cDNA library was constructed from mRNA of excised immature galls of resistant cotton (M-249) roots at 10 DAI with RKN [*Meloidogyne incognita* (Kofoid and White) Chitwood, race 3]. Using the gene-specific primer R1 and λZAPII vector primer T3, a DNA fragment was amplified by polymerase chain reaction (PCR) from the cDNA library stock culture. This fragment was cloned into pGEM-T Easy vector (Promega) and sequenced with ABI PRISM 310 DNA Genetic Analyzer (Perkin-Elmer). Based on the identified 5′ end cDNA sequence, a forward primer, F1 (5′-TTAAAAATGGCTTGTCCTCCAACTCATAAAGT-3′, nt positions 16–47 in Fig. 1) was designed and then used together with the vector primer T7 to amplify the 3′ end cDNA sequence by PCR. The amplified 3′ cDNA fragment was cloned and then sequenced as described. Another reverse primer, R2 (5′-AATTGCAACCGCTCC ACATGATTTGTGCT-3′, nt positions 493–522 in Fig. 1), was designed based on the identified 3′ end cDNA sequences and paired with T3 primer for PCR amplification to further confirm the 5′ end cDNA sequence. The two primers, F1 and R2, were used to amplify the genomic DNA sequences from the resistant line. In all PCR reactions, *Pfu* DNA polymerase (Stratagene) possessing the 3′ to 5′ proofreading proofreading exonuclease activity was used for accurate sequence determination.

The final assembled nucleotide sequence for *MIC-3* (*Meloidogyne* induced cotton-gene) has an open reading frame of 423 bp with the start codon at nt position 22, which is in the context of AAAATGG that fulfills the criterion of a Kozak sequence (A/GNNATGG) [13] (Fig. 1). The *MIC-3* gene encodes a putative protein of 141 amino acids with an estimated molecular weight of 15,357 and isoelectric point of 5.5. A single intron of 77 bp contains the conserved GT and AG sequences as the 5′ and 3′ splicing sites, respectively. Seven *MIC-3* transcripts, differing only in length at the 3′ regions, and a poly(A) sequence of at least 20 nt were amplified from the cDNA library by primers F1 and T7. The alternative polyadenylation sites are located at nt positions 679, 681, 713, 715, 716, 718, and 803 (Fig. 1). The predominant transcripts were the longest ones present in more than half of the amplified cDNAs. Three putative polyadenylation signals (AAUAUA, AUUAAA, AAUAAA) [14] are at nt positions 651, 688, and 782. The canonical signal AAUAAA, which is processed more efficiently than the variant signals AAUAUA and AUUAAA, is generally located at the 3′-most terminal site, and the predominant form of alternatively polyadenylated transcripts is normally the longest one [14]. The two alternative polyadenylation

```
GAACTAAAAAAGGAGTTAAAAATGGCTTGTCCTCCAACTCATAAAGTGGTAAAAGGTGAG   60
                    M  A  C  P  P  T  H  K  V  V  K  G  E    13
CTTAAGAATGGTTCTGATGACTGCCTTGCGAGCTTGATGTATTCGAATTACCAAGGTGCT  120
 L  K  N  G  S  D  D  C  L  A  S  L  M  Y  S  N  Y  Q  G  A    33
CCGGTAACGATTAAAAAAGATTCATTACCCACTCCATTTGCGCAGAATATGATGGTCGAC  180
 P  V  T  I  K  K  D  S  L  P  T  P  F  A  Q  N  M  M  V  D    53
GGTCTGTACGGAGGCCTGGTGTATGATGTTGTTAGGGTTAAGTGGATTATTCTTTGGACC  240
 G  L  Y  G  G  L  V  Y  D  V  V  R  V  K  W  I  I  L  W  T    73
ACTGATTGTAAGgtattaattaatcaatatatttttcatttattttgattttgtagcaga  300
 T  D  C  K                                                    77
agaaattaattaattaaactatatatcagGTGGCTACTAAGATCATTCCGACCGAAAACC  360
                             V  A  T  K  I  I  P  T  E  N     87
ATATTGTTTGGGAGGATATAGTGAGCATCCTTCAGCCCTACAATAGCTCTGACATCTTGC  420
 H  I  V  W  E  D  I  V  S  I  L  Q  P  Y  N  S  S  D  I  L   107
CCTTATCATGTGGGGGTGCATTTTCCTCCGAAGCTCATATCCATGCAAATGCAGATGGAT  480
 P  L  S  C  G  G  A  F  S  S  E  A  H  I  H  A  N  A  D  G   126
CTCTAAACCTGACAGCACAAATCATGTGGAGCGGTTGCAATTAACCTGCCTCTCTTTGCA  540
 S  L  N  L  T  A  Q  I  M  W  S  G  C  N  *                  141
TCATATATTGCTCCATGCTTATCTCTATCGGTTTATGCTGTGTGACCCATGGAGGCTGTC  600

TACTTCTTTATATTAATAATACTACTAAATAAATTTGACAGCCACCCCATAATATACTAC  660

CGGATTCCTACTTGTTCTACCTTTCCTATTAAACATTGTGTGCCTTGTTTCTCCTCTTTC  720
                          Δ  Δ                        Δ ΔΔ Δ
TCTTCTTTTTTTTTTTTTTATCTTTTCTGTAATAATATATATACCCACGATTGTGTGAAAT  780

TAATAAAAAAATTCTATGTCTAT                                       803
                      Δ
```

Fig. 1. Nucleotide and derived amino acid sequences of the cotton *MIC-3* gene. The sequence of the 77 bp intron is shown in lowercase letters, and the conserved intron splicing sequences GT and AG are underlined. The stop codon is indicated by an asterisk. The three putative polyadenylation signals are double-underlined and the seven alternative polyadenylation sites are indicated by open arrowheads.

sites at nt positions 679 and 681 are 28 and 30 nt downstream of the AAUAUA signal, and the four alternative polyadenylation sites at nt positions 713, 715, 716, and 718 are 25, 27, 28, and 30 nt downstream of the AUUAAA signal. The major polyadenylation site at nt position 803 is 21 nt downstream of the AAUAAA signal. The percentage of genes whose mRNA transcripts use two or more polyadenylation cleavage sites downstream from a single polyadenylation signal is considerably high in mammals, with 44% in human, 22% in mouse, and 22% in rat [27]. Alternative polyadenylation is frequently associated with tissue- or disease-specific expression [28,29]. Interestingly, the *MIC-3* gene and predicted protein showed no significant homology to known sequences in any of the available databases.

Genomic Southern blot analysis using ^{32}P-labeled *MIC-3* cDNA probe revealed at least six hybridizing bands dependent on the restriction enzyme employed, suggesting that *MIC-3* gene belonged to a multi-gene family (Fig. 2). Northern analysis was used to determine the expression pattern of the *MIC-3* gene. Seedlings of RKN-resistant (M-249) and recurrent susceptible parent (ST213) were inoculated with J2 RKN [9]. The small immature galls at 10 DAI were excised from roots of both resistant and susceptible plants [12]. Equivalent sections of roots were also excised from non-inoculated controls under the same conditions.
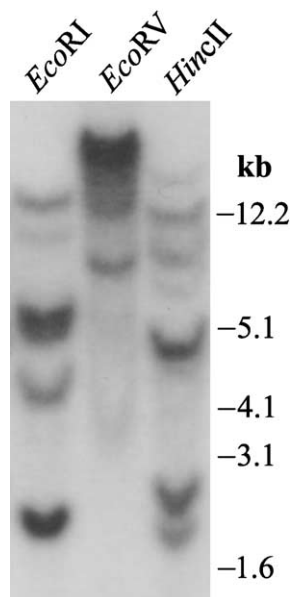

Fig. 2. Genomic Southern blot analysis of *MIC-3* gene family in cotton. Genomic DNA (10 μg) extracted from young leaves of M-249 (RKN-resistant isoline) was digested with restriction enzymes *Eco*RI, *Eco*RV, and *Hinc*II, respectively, separated by agarose gel electrophoresis, blotted, and hybridized with ^{32}P-labeled *MIC-3* cDNA probe. The membrane was washed under high-stringency (twice for 10 min at room temperature in 1 × SSC plus 0.1% SDS, once for 50 min at 65 °C in 1 × SSC plus 0.1% SDS, and once for 20 min at 65 °C in 0.2 × SSC plus 0.1% SDS) and visualized by autoradiography at − 80 °C in the presence of an intensifying screen.
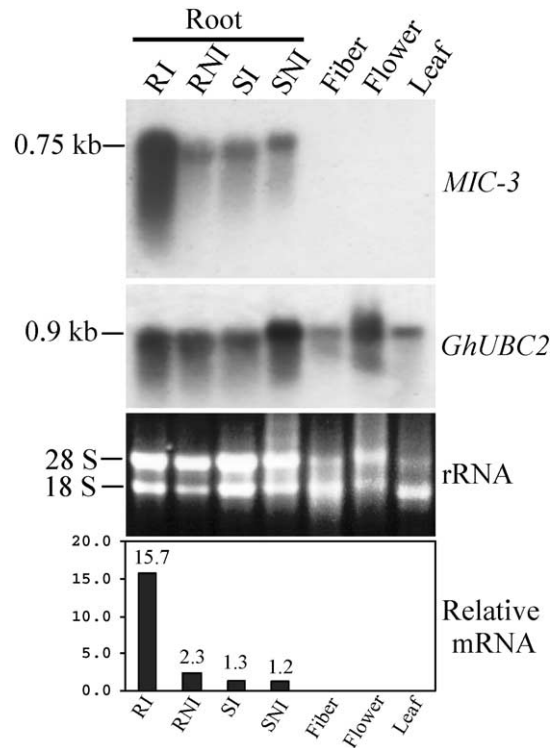

Fig. 3. Northern blot analysis of the expression of the cotton *MIC-3* gene. Total RNA samples (10 μg each) isolated from excised, immature galls of resistant (M-249) and recurrent, susceptible parent (ST213) plants at 10 DAI with RKN (RI and SI, respectively), equivalent sections of roots of resistant and susceptible non-inoculated plants (RNI and SNI, respectively), 15 DPA fibers, yellow flowers, and young leaves were separated by agarose gel electrophoresis, blotted, and hybridized with ^{32}P-labeled *MIC-3* cDNA probe. Following hybridization, the membrane was washed under high-stringency (twice for 10 min at room temperature in 1 × SSPE plus 0.1% SDS and twice for 20 min at 65 °C in 0.5 × SSPE plus 0.1% SDS) and exposed to X-ray film at − 80 °C in the presence of an intensifying screen. The relative mRNA levels of *MIC-3* were determined by the ratio of the intensity of the 0.75-kb *MIC-3* mRNA to the EtBr-stained 28S and 18S rRNA bands using Scion Image for Windows (Scion Corporation, http://www.scioncorp.com). After removal of the *MIC-3* probe, the blot was rehybridized with a control probe *GhUBC2* (cotton class I ubiquitin-conjugating enzyme gene) known to be expressed in all cotton tissues [15].

Aliquots of 10 μg total RNA isolated from the root samples, 15 DPA (days post-anthesis) fibers, yellow flowers, and young leaves were fractionated by gel electrophoresis, blotted, hybridized with a ^{32}P-labeled *MIC-3* cDNA probe, and washed under high stringency conditions. A 0.75-kb transcript was detected in the root samples but not in fibers, flowers, or leaves; the transcript level of *MIC-3* was at least six times higher in the immature root galls of the resistant plants (Fig. 3). The estimated size (0.75 kb) of the *MIC-3* transcript was very close to the length of its longest cDNA sequence of 747 bp (including the 21 nt poly-A sequence), indicating that the cDNA sequence reported was full-length. Rehybridization of the same blot with a control probe *GhUBC2* (cotton class I ubiquitin-conjugating enzyme gene) known to be expressed in all tissues [15], verified

the observed root-specific and induced expression of the *MIC-3* gene in the resistant inoculated plants (Fig. 3).

The expression pattern for the *MIC-3* gene, size of the predicted protein, and calculated p*I* were consistent with our previous data on expression of the 14-kDa protein [12]. These results suggest that the *MIC-3* gene may encode the 14-kDa protein. The deduced *MIC-3* amino acid sequence includes (MMVDGLYG) (50–57) (Fig. 1) homologous to the short internal sequence (MKVDGYYG) obtained from the gel-purified 14 kDa protein [12]. We suspect that the two amino acid differences reflect less than homogeneous purification of the internal fragment of the 14 kDa, leading to some inaccuracy in the protein sequencing. The *MIC-3* protein has been successfully expressed in *E. coli* at the expected size (data not shown) and will be used as antigen to develop antiserum to the protein. Western blot analysis of the cotton 14 kDa protein using the *MIC-3* antiserum will further verify their relationship.

To our knowledge, this is the first report of a root-specific gene with increased expression specifically in nematode-resistant plants following nematode infection. Eight cDNAs have been identified with increased expression in both resistant and susceptible tomato after RKN inoculation, suggesting that they are likely to play a role in the more general, root-defense system [11]. Many genes which are up-regulated in nematode-induced feeding cells have been isolated [16–25], and their elevated expression is likely to be important for the establishment and maintenance of the nematode-induced feeding structures or for the common defense response in susceptible plants [11,16]. Tomato RKN resistance is conferred by a single, dominant gene, *Mi-1.2*, the only cloned RKN resistance gene in plants [3]. The signal transduction pathway leading to localized cell death may be initiated by the interaction between *Mi-1.2* and a plant protein induced by the nematode infection [26]. As a gene whose expression is specifically enhanced in RKN-induced, immature galls of resistant cotton line M-249, *MIC-3* is likely involved in the resistance mechanism mediated by the cotton RKN resistance genes. Characterization of the *MIC* gene family members and their expression in other cotton backgrounds with the same source of resistance genes will be needed to further substantiate a role in the resistance mechanism. The over-expression of *MIC-3* gene in RKN susceptible cotton roots and the reduction or elimination of *MIC-3* product by antisense mRNA expression in resistant plants will more directly test the hypothesis.

Acknowledgements

References

[1] J.N. Sasser, Root-knot nematodes: a global menace to crop production, Plant Dis. 64 (1980) 36–41.

[2] V.M. Williamson, R.S. Hussey, Nematode pathogenesis and resistance in plants, Plant Cell 8 (1996) 1735–1745.

[3] S. Milligan, J. Bodeau, J. Yaghoobi, I. Kaloshian, P. Zabel, V.M. Williamson, The root-knot nematode resistance gene *Mi* from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes, Plant Cell 10 (1998) 1307–1319.

[4] R.L. Shepherd, Registration of Auburn 623RNR cotton germplasm, Crop Sci. 14 (1974) 911.

[5] R.L. Shepherd, Registration of three germplasm lines of cotton, Crop Sci. 22 (1982) 692.

[6] R.L. Shepherd, Registration of three root-knot resistant cotton germplasm lines, Crop Sci. 27 (1987) 153.

[7] R.L. Shepherd, Trangressive segregation for root-knot nematode resistance in cotton, Crop Sci. 14 (1974) 872–875.

[8] R.L. Shepherd, M.G. Huck, Progression of root-knot nematode symptoms and infection on resistant and susceptible cottons, J. Nematol. 21 (1989) 235–241.

[9] R.G. Creech, J.N. Jenkins, B. Tang, G.W. Lawrence, J.C. McCarty, Cotton resistance to root-knot nematode: I. Penetration and reproduction, Crop Sci. 35 (1995) 365–368.

[10] J.N. Jenkins, R.G. Creech, B. Tang, G.W. Lawrence, J.C. McCarty, Cotton resistance to root-knot nematode: II. Post penetration development, Crop Sci. 35 (1995) 369–373.

[11] K.N. Lambert, B.J. Ferrie, G. Nombela, E.D. Brenner, V.M. Williamson, Identification of genes whose transcripts accumulate rapidly in tomato after root-knot nematode infection, Physiol. Mol. Plant Pathol. 55 (1999) 341–348.

[12] F.E. Callahan, J.N. Jenkins, R.G. Creech, G.W. Lawrence, Changes in cotton root protein correlated with resistance to root knot nematode development, J. Cotton Sci. 1 (1997) 38–47.

[13] M. Kozak, An analysis of vertebrate mRNA sequences: intimations of translational control, J. Cell Biol. 115 (1991) 887–903.

[14] E. Beaudoing, S. Freier, J.R. Wyatt, J.M. Claverie, D. Gautheret, Patterns of variant polyadenylation signal usage in human genes, Genome Res. 10 (2000) 1001–1010.

[15] X.D. Zhang, Isolation and characterization of two cotton cDNAs and corresponding genes encoding ubiquitin-conjugating enzymes. MS Thesis, Mississippi State University, 1998.

[16] C.H. Opperman, C.G. Taylor, M.A. Conkling, Root-knot nematode-directed expression of a plant root-specific gene, Science 263 (1994) 221–223.

[17] D.M. Bird, M.A. Wilson, DNA sequence and expression analysis of root-knot nematode-elicited giant cell transcripts, Mol. Plant-Microbe Interact. 7 (1994) 419–424.

[18] D.M. Bird, Manipulation of host gene expression by root-knot nematodes, J. Parasitol. 82 (1996) 881–888.

[19] N. Barthels, F.M. van der Lee, J. Klap, O.J. Goddijn, M. Karimi, P. Puzio, F.M. Grundler, S.A. Ohl, K. Lindsey, L. Robertson, W.M. Robertson, M. Van Montagu, G. Gheysen, P.C. Sijmons, Regulatory sequences of *Arabidopsis* drive reporter gene expression in nematode feeding structures, Plant Cell 9 (1997) 2119–2134.

[20] D. Hermsmerier, M. Mazarei, T.J. Baum, Differential display analysis of the early compatible interaction between soybean and the soybean cyst nematode, Mol. Plant-Microbe Interact. 11 (1998) 1258–1263.

[21] A. Niebel, J. de Almeida Engler, A. Hemerly, P. Ferreira, D. Inze, M. Van Montagu, G. Gheysen, Induction of cdc2a and cyc1At expression in *Arabidopsis thaliana* during early phases of nematode-induced feeding cell formation, Plant J. 10 (1996) 1037–1043.

[22] G. Strittmatter, G. Gheysen, V. Gianinazzi-Pearson, K. Hahn, A. Niebel, W. Rohde, E. Tacke, Infections with various types of organisms stimulate transcription from a short promoter fragment of the potato gst1 gene, Mol. Plant-Microbe Interact. 9 (1996) 68–73.

[23] A. Niebel, K. Heungens, N. Barthels, D. Inze, M. Van Montagu, G. Gheysen, Characterization of a pathogen-induced potato catalase and its systemic expression upon nematode and bacterial infection, Mol. Plant-Microbe Interact. 8 (1995) 371–378.

[24] T.C. Vrain, Engineering natural and synthetic resistance for nematode management, J. Nematol. 31 (1999) 424–436.

[25] W. Van der Eycken, J. de Almeida Engler, D. Inze, M. Van Montagu, G. Gheysen, A molecular study of root-knot nematode-induced feeding sites, Plant J. 9 (1996) 45–54.

[26] C.F. Hwang, A.V. Bhakta, G.M. Truesdell, W.M. Pudlo, V.M. Williamson, Evidence for a role of the N terminus and leucine-rich repeat region of the *Mi* gene product in regulation of localized cell death, Plant Cell 12 (1998) 1319–1329.

[27] E. Pauws, A.H. van Kampen, S.A. van de Graaf, J.J. de Vijlder, C. Ris-Stalpers, Heterogeneity in polyadenylation cleavage sites in mammalian mRNA sequences: implications for SAGE analysis, Nucleic Acids Res. 29 (2001) 1690–1694.

[28] G. Edwalds-Gilbert, K.L. Veraldi, C. Milcarek, Alternative poly(A) site selection in complex transcription units: means to an end? Nucleic Acids Res. 25 (1997) 2547–2561.

[29] E. Beaudoing, D. Gautheret, Identification of alternate polyadenylation sites and analysis of their tissue distribution using EST data, Genome Res. 11 (2001) 1520–1526.